Estimating the conditional tail index with an integrated conditional log-quantile estimator in the random covariate case

Laurent Gardes, Gilles Stupfler

Estimating the conditional tail index with an integrated conditional log-quantile estimator in the random covariate case

Laurent Gardes[(1)] & Gilles Stupfler[(2)]

[(1)] Université de Strasbourg & CNRS, IRMA, UMR 7501, 7 rue René Descartes, 67084 Strasbourg Cedex, France
[(2)] Aix Marseille Université, CNRS, EHESS, Centrale Marseille, GREQAM UMR 7316, 13002 Marseille, France

Abstract. It is well known that the tail behavior of a heavy-tailed distribution is controlled by a parameter called the tail index. Such a parameter is therefore of primary interest in extreme value analysis, particularly to estimate extreme quantiles. In various applications, the random variable of interest can be linked to a finite-dimensional random covariate. In such a situation, the tail index is function of the covariate and is referred to as the conditional tail index. The goal of this paper is to provide a class of estimators of this quantity. The pointwise weak consistency and asymptotic normality of these estimators are established. We illustrate the finite sample performance of our technique on a simulation study and on a real hurricane data set.

AMS Subject Classifications: 62G05, 62G20, 62G30, 62G32.

Keywords: Heavy-tailed distribution, tail index, random covariate, consistency, asymptotic normality.

1 Introduction

Studying extreme events is relevant in numerous fields of statistical applications. In hydrology for example, it is of interest to estimate the maximum level reached by seawater along a coast over a given period, or to study extreme rainfall at a given location; in actuarial science, a major problem for an insurance firm is to estimate the probability that a claim so large that it represents a threat to its solvency is filed. A particular branch of extreme value analysis focuses on the study of heavy-tailed random variables, that is, those random variables whose distribution function F is such that, for all $\lambda > 0$, $(1 - F(\lambda x))/(1 - F(x)) \to \lambda^{-1/\gamma}$ as x goes to infinity, where $\gamma > 0$ is the so-called tail index. The parameter γ drives the asymptotic behavior of F in its right tail, which makes its estimation necessary if we are interested in the extremes of the associated random variable. The estimation of the tail index has therefore been extensively studied in the literature. Recent overviews on univariate tail index estimation can be found in Beirlant *et al.* [2] and de Haan and Ferreira [22].

In practical applications, the variable of interest Y can often be linked to a covariate X. For instance, the value of rainfall at a given location depends on its geographical coordinates; in actuarial science, the claim size depends on the sum insured by the policy. In this situation, the tail index of the random variable Y given $X = x$ is a function of x to which we shall refer as the conditional tail index. Its estimation has first been considered in the "fixed design" case, namely when the covariates are nonrandom. Smith [30] and Davison and Smith [12] considered a regression model while Hall and Tajvidi [23] used a semi-parametric approach to estimate the conditional tail index. Fully non-parametric methods have been developed using splines (see Chavez-Demoulin and Davison [7]), local polynomials (see Davison and Ramesh [11]), a moving window approach (see Gardes and Girard [15]), a nearest neighbor approach (see Gardes and Girard [16]), and a conditional quantile-based technique (see Gardes *et al.* [18]), among others.

Despite the great interest in practice, the study of the random covariate case has been initiated only recently. We refer to the works of Wang and Tsai [32], based on a maximum likelihood approach, Daouia *et al.* [9] who used a fixed number of non parametric conditional quantile estimators to estimate the conditional tail index, later generalized in Daouia *et al.* [10] to a regression context with conditional response distributions belonging to the general max-domain of attraction, Gardes and Girard [17] who introduced a local generalized Pickands-type estimator (see Pickands [27]), Goegebeur *et al.* [20] who studied a nonparametric regression estimator whose strong uniform properties are examined in Goegebeur *et al.* [21], Stupfler [31] who introduced a generalization of the popular moment estimator of Dekkers *et al.* [13] and Gardes and Stupfler [19] who worked on a smoothed local Hill estimator (see Hill [24]) related to the work of Resnick and Stărică [28].

The aim of this paper is to introduce an estimator of the conditional tail index based on the integration of a conditional log-quantile estimator. This type of estimators is similar to the one of Gardes and Girard [15]; our aim is to prove its consistency and asymptotic normality when the covariates are random, as well as to examine its applicability on numerical examples and on real data. Our paper is organized as follows: we define our conditional tail index estimator in Section 2, its asymptotic properties are stated in Section 3, a simulation study is provided in Section 4 and we showcase our estimator on a set of real hurricane data in Section 5. We offer a couple of concluding remarks in Section 6. All the auxiliary results and proofs are deferred to the Appendix.

2 Framework

We let $(X_1, Y_1), \dots, (X_n, Y_n)$ be n independent copies of a random pair $(X, Y) \in \mathcal{E} \times \mathbb{R}_+$, where $(\mathcal{E}, d)$ is a metric space. We assume that for any $x \in \mathcal{E}$, the conditional distribution function $y \mapsto F(y|x) := \mathbb{P}(Y \leq y | X = x)$ of Y given $X = x$ belongs to the set $\mathcal{RV}_{-1/\gamma(x)}$ of regularly varying functions (at infinity) of index $-1/\gamma(x) < 0$. Recall that a function $G \in \mathcal{RV}_a$, $a \in \mathbb{R}$ if G is nonnegative and for all $\lambda > 0$, $G(\lambda y)/G(y) \to \lambda^a$ as y goes to infinity. This is the adaptation of the standard extreme-value framework to the case when there is a covariate. An equivalent assumption (see Bingham *et al.* [5, Proposition 1.5.15]) is:

(M1) For any $x \in \mathcal{E}$, the conditional quantile function $\alpha \mapsto q(\alpha|x) := F^{\leftarrow}(1 - \alpha|x) = \inf\{y \in \mathbb{R} \mid F(y|x) \geq 1 - \alpha\} \in \mathcal{RV}_{-\gamma(x)}$.

Our goal is to estimate the conditional tail index γ at a point $x \in \mathcal{E}$. Remark first that, under **(M1)**, for $u \in (0, 1)$ small enough and $\alpha \in (0, u)$, $\log q(\alpha|x)/q(u|x) \approx \gamma(x) \log(u/\alpha)$. Hence, for any measurable function $\Psi(.|x, u)$ on $(0, u)$ such that

$$\int_0^u \Psi(\alpha|x, u) \log(u/\alpha)\, d\alpha = 1, \tag{1}$$

one has

$$\int_0^u \Psi(\alpha|x, u) \log \frac{q(\alpha|x)}{q(u|x)} d\alpha \approx \gamma(x). \tag{2}$$

We propose to estimate $\gamma(x)$ by replacing in the previous approximation the conditional quantile function $q(.|x)$ by a consistent estimator of this quantity. To this end, let $\mathbb{I}\{.\}$ denote the indicator function and, for any $h > 0$, $B(x, h) := \{x' \in \mathcal{E} \mid d(x, x') \leq h\}$ denote the closed ball in $\mathcal{E}$ with center x and radius h. The total number of covariates belonging to the ball $B(x, h)$ is given by

$$M(x, h) = \sum_{i=1}^n \mathbb{I}\{X_i \in B(x, h)\}.$$

The conditional distribution function $F(.|x)$ is estimated by:

$$\widehat{F}_n(y|x, h_x) = \frac{1}{M(x, h_x)} \sum_{i=1}^n \mathbb{I}\{Y_i \leq y\} \mathbb{I}\{X_i \in B(x, h_x)\},$$

where $h_x = h_x(n)$ is a positive sequence converging to 0. The associated estimator of the conditional quantile function $q(.|x)$ is then, for $\alpha \in (0, 1)$,

$$\widehat{q}_n(\alpha|x, h_x) = \widehat{F}_n^{\leftarrow}(1 - \alpha|x, h_x) = \inf\{y \in \mathbb{R} \mid \widehat{F}_n(y|x, h_x) \geq 1 - \alpha\}.$$

Replacing $q(.|x)$ by $\widehat{q}_n(.|x, h_x)$ in (2), our class of estimators of $\gamma(x)$ is given for a $(0, 1)$-valued measurable function u_x converging to 0 at infinity by:

$$\widehat{\gamma}(x, u_x, h_x) = \int_0^{U_x} \Psi(\alpha|x, U_x) \log \frac{\widehat{q}_n(\alpha|x, h_x)}{\widehat{q}_n(U_x|x, h_x)} d\alpha, \tag{3}$$

in which $U_x = u_x(M(x, h_x))$ and $\Psi(.|x, u)$ is an integrable function on $(0, u)$ satisfying (1). The estimator $\widehat{\gamma}(x, u_x, h_x)$ is thus a weighted integral of an estimator of the conditional log-quantile function.
We conclude this section by pointing out that particular choices of the function $\Psi(.|x, u)$ actually yield generalizations of some well-known tail index estimators to the conditional framework. Let $k_x := U_x M(x, h_x)$. The choice $\Psi(.|x, u) = u^{-1}$ yields:

$$\widehat{\gamma}^H(x, u_x, h_x) = \frac{1}{k_x} \sum_{i=1}^{\lfloor k_x \rfloor} \log \frac{\widehat{q}_n((i-1)/M(x, h_x)|x, h_x)}{\widehat{q}_n(k_x/M(x, h_x)|x, h_x)}, \tag{4}$$

which is the straightforward adaptation of the classical Hill estimator (see Hill [24]). Similarly, letting $\Psi(.|x, u) = u^{-1}(\log(u/.) - 1)$ entails, after some algebra:

$$\widehat{\gamma}^Z(x, u_x, h_x) = \frac{1}{k_x} \sum_{i=1}^{\lfloor k_x \rfloor} \log\left(\frac{k_x}{i}\right) \left\{ i \log \frac{\widehat{q}_n((i-1)/M(x, h_x)|x, h_x)}{\widehat{q}_n(i/M(x, h_x)|x, h_x)} \right\}.$$

This estimator can be seen as a generalization of the Zipf estimator (see Kratz and Resnick [26], Schultze and Steinebach [29]).

3 Asymptotic properties

3.1 Main results

We start by stating the weak consistency of the estimator (3). To this end, an additional hypothesis is required.

(A1) The function $\Psi(.|x, u)$ satisfies:

$$\limsup_{u \downarrow 0} \int_0^u |\Psi(\alpha|x, u)| d\alpha < \infty,$$

and for all $u \in (0, 1)$ and $\beta \in (0, u]$,

$$\frac{u}{\beta} \int_0^\beta \Psi(\alpha|x, u) d\alpha = \Phi(\beta/u|x),$$

where $\Phi(.|x)$ is a square-integrable nonincreasing probability density function on $(0, 1)$.

Note that condition **(A1)** is satisfied by the two functions $\Psi(.|x, u) = u^{-1}$ and $\Psi(.|x, u) = u^{-1}(\log(u/.) - 1)$ with $\Phi(.|x) = 1$ and $\Phi(.|x) = -\log(.)$ respectively. We also assume in all what follows that $q(.|x)$ is continuous and decreasing. Particular consequences of this condition include that $F(q(\alpha|x)|x) = 1 - \alpha$ for any $\alpha \in (0, 1)$ and that given $X = x$, Y has an absolutely continuous distribution with probability density function $f(.|x)$. For $0 < \alpha_1 < \alpha_2 < 1$, we finally introduce the quantity:

$$\omega(\alpha_1, \alpha_2, x, h_x) = \sup_{\alpha \in [\alpha_1, \alpha_2]} \sup_{x' \in B(x, h_x)} \left| \log \frac{q(\alpha|x')}{q(\alpha|x)} \right|,$$

which is the uniform oscillation of the log-quantile function in its second argument. Such a quantity is also studied in Gardes and Stupfler [19], for instance. Letting $m_x(h_x) = n\mathbb{P}(X \in B(x, h_x))$ be the average number of covariates which belong to $B(x, h_x)$, the weak consistency of our family of estimators is established in the following theorem.

Theorem 1. *Assume that conditions* **(M1)** *and* **(A1)** *are satisfied. Assume further that $m_x(h_x) \to \infty$ as $n \to \infty$ and that $u_x \in \mathcal{RV}_{-a(x)}$ with $a(x) \in (0, 1)$. If, for some $\delta > 0$,*

$$\omega\left([m_x(h_x)]^{-1-\delta}, 1 - [m_x(h_x)]^{-1-\delta}, x, h_x\right) \to 0, \tag{5}$$

then it holds that $\widehat{\gamma}(x, u_x, h_x) \xrightarrow{\mathbb{P}} \gamma(x)$ as $n \to \infty$.

Note that $u_x(m_x(h))m_x(h) \to \infty$ is the average number of observations used to compute our estimator of $\gamma(x)$. The conditions in Theorem 1 are thus analogues of the classical hypotheses in the estimation of the tail index. Besides, condition (5) ensures that the distribution of Y given $X = x'$ is close enough to that of Y given $X = x$ when x' is in a sufficiently small neighborhood of x.

Our aim is now to establish an asymptotic normality result. First, recall that under **(M1)**, the conditional quantile function may be written as follows:

$$\forall t > 1, \; q(t^{-1}|x) = c(t|x) \exp\left(\int_1^t \frac{\Delta(v|x) - \gamma(x)}{v} dv \right),$$

where $c(.|x)$ is a positive function converging to a positive constant at infinity and $\Delta(.|x)$ is a measurable function converging to 0 at infinity, see Bingham *et al.* [5, Theorem 1.3.1]. We introduce the following classical second-order condition:

(M2) Condition **(M1)** holds, $c(.|x)$ is a constant function equal to $c(x) > 0$, the function $\Delta(.|x)$ has ultimately constant sign at infinity and $|\Delta(.|x)| \in \mathcal{RV}_{\rho(x)}$, with $\rho(x) < 0$.

In condition **(M2)**, $\rho(x)$ is called the conditional second-order parameter of the distribution. This condition is commonly used when studying tail index estimators and makes it possible to control the asymptotic bias of the estimator $\widehat{\gamma}(x, u_x, h_x)$. We also introduce a further assumption on the weighting function $\Phi(.|x)$, which is similar in spirit to a condition introduced in Beirlant *et al.* [1]. To write down this condition, we note that if **(A1)** holds then

$$\forall \beta \in (0, 1), \; 0 \leq \beta \Phi(\beta|x) \leq \int_0^{\beta/2} |\Psi(\alpha|x, 1/2)| d\alpha$$

and the right-hand side converges to 0 as $\beta \downarrow 0$, so that we may extend the definition of the map $t \mapsto t\Phi(t|x)$ by saying it is 0 at $t = 0$.

(A2) Condition **(A1)** holds, there is $\kappa > 0$ such that $\Phi^{2+\kappa}(.|x)$ is integrable on $(0,1)$ and there exists a positive function $g(.|x)$, which is either continuous on $[0,1]$ or nonincreasing on $(0,1)$, such that for any $k > 1$ and $i \in [1,k)$,

$$|i\Phi\left(i/k|x\right) - (i-1)\Phi\left((i-1)/k|x\right)| \leq g\left(i/k|x\right),$$

where the function $g(.|x)\max(\log(1/.),1)$ is integrable on $(0,1)$.

Note that condition **(A2)** is satisfied for instance by the functions $\Psi(.|x,u) = u^{-1}$ and $\Psi(.|x,u) = u^{-1}(\log(u/.) - 1)$ mentioned at the end of Section 2 with $g(.|x) = 1$ for the first one and, for the second one, $g(.|x) = -\log(.) + 1$. Our asymptotic normality result is the following:

Theorem 2. *Assume that conditions* **(M2)** *and* **(A2)** *are satisfied. Assume further that $m_x(h_x) \to \infty$ as $n \to \infty$, that $u_x \in \mathcal{RV}_{-a(x)}$ with $a(x) \in (0,1)$ and $(zu_x(z))^{1/2}\Delta(1/u_x(z)|x) \to \lambda(x) \in \mathbb{R}$ as $z \to \infty$. If for some $\delta > 0$,*

$$v_x^{1/2}\omega\left([m_x(h_x)]^{-1-\delta}, 1 - [m_x(h_x)]^{-1-\delta}, x, h_x\right) \to 0 \tag{6}$$

where $v_x = m_x(h_x)u_x(m_x(h_x))$, then it holds that

$$v_x^{1/2}(\widehat{\gamma}(x,u_x,h_x) - \gamma(x)) \xrightarrow{d} \mathcal{N}\left(\lambda(x)\mathcal{AB}_x(\Phi,\rho(x)), \gamma^2(x)\mathcal{AV}_x(\Phi)\right)$$

as $n \to \infty$, with

$$\mathcal{AB}_x(\Phi,\rho(x)) = \int_0^1 \Phi(\alpha|x)\alpha^{-\rho(x)}d\alpha \quad \text{and} \quad \mathcal{AV}_x(\Phi) = \int_0^1 \Phi^2(\alpha|x)d\alpha.$$

Our asymptotic normality result thus holds under generalizations of the common hypotheses on the model and on u_x and h_x, provided the conditional distributions of Y at two neighboring points are sufficiently close.
We conclude this paragraph by noting that these results are similar in spirit to results obtained in the literature for other conditional tail index or conditional extreme-value index estimators, see *e.g.* Gardes and Stupfler [19] and Stupfler [31]. The main disadvantage of formulating the hypotheses in terms of the uniform oscillation ω is that they cannot immediately be translated in terms of conditions on u_x and h_x. In our next paragraph, we give alternative, simple conditions for our main results to hold.

3.2 Discussion of the hypotheses

As a starting point, we note that if X has a probability density function f with respect to the Lebesgue measure on $\mathcal{E} = \mathbb{R}^d$ equipped with the Euclidean norm $\|.\|$ then sufficient conditions for $m_x(h_x) \to \infty$ are that $h_x \to 0$, $nh_x^d \to \infty$,

$f(x)>0$ and f is continuous at x. Indeed, in this case, if $\mathcal{V}$ denotes the volume of the unit ball of $\mathbb{R}^d$, a change of variables entails:

$$m_x(h_x) = n\int_{B(x,h_x)} f(s)ds = nh_x^d f(x)\left(\mathcal{V} + \int_{\|v\|\le 1}\left[\frac{f(x+h_xv)}{f(x)} - 1\right]dv\right).$$

Since f is continuous at x, we get $m_x(h_x) = nh_x^d\mathcal{V}f(x)(1+\mathrm{o}(1)) \to \infty$. Furthermore, we point out that if the functions γ, $\log c(t|.)$ and $\Delta(t|.)$ satisfy a Hölder condition, namely:

$$\begin{aligned}
\sup_{x'\in B(x,h_x)} |\gamma(x') - \gamma(x)| &= \mathrm{O}(h_x^\beta), \\
\sup_{t^{-1}\in K_{x,\delta}(h_x)}\ \sup_{x'\in B(x,h_x)} |\log c(t|x') - \log c(t|x)| &= \mathrm{O}(h_x^\beta) \\
\text{and}\ \sup_{t^{-1}\in K_{x,\delta}(h_x)}\ \sup_{x'\in B(x,h_x)} |\Delta(t|x') - \Delta(t|x)| &= \mathrm{O}(h_x^\beta),
\end{aligned}$$

where $\beta>0$ and $K_{x,\delta}(h_x)$ is the interval $[(m_x(h_x))^{-1-\delta}, 1-(m_x(h_x))^{-1-\delta}]$, then (5) is a consequence of the convergence $h_x^\beta \log m_x(h_x) \to 0$. In the aforementioned context when X has a probability density function, this condition becomes $h_x^\beta \log n \to 0$ as $n\to\infty$. Such conditions were already considered in Stupfler [31].

As an illustration, we now compute the optimal rate of convergence of our estimator when $\mathcal{E} = \mathbb{R}^d$ and X has a probability density function. Let $a(x)\in(0,1)$ and $b(x)\in(0,1/d)$. We take $\log(h_x) = -b(x)\log(n)$ and $\log(nu_x(n)) = (1-a(x))\log(n)$. In this context, the rate of convergence of the estimator is essentially $(m_x(h_x)u_x(m_x(h_x)))^{1/2} = n^{(1-db(x))(1-a(x))/2}$. Besides, since $\Delta(.|x)$ is regularly varying with index $\rho(x)<0$, the conditions for Theorem 2 to hold are then essentially:

$$1-a(x)+2a(x)\rho(x)\le 0 \quad\text{and}\quad 1-a(x)-2\beta b(x)\le 0.$$

The problem thus amounts to maximizing the function $(a,b)\mapsto(1-db)(1-a)$ under these conditions. The solution is:

$$(a^*(x), b^*(x)) = \left(\frac{1}{1-2\rho(x)},\ \frac{\rho(x)}{d\rho(x)+\beta(2\rho(x)-1)}\right),$$

which yields the optimal rate of convergence $n^{\beta\rho(x)/(d\rho(x)+\beta(2\rho(x)-1))}$. Note that setting $d=0$, *i.e.* considering the case when there is no covariate, we recover the optimal rate of convergence of the Hill estimator, see *e.g.* de Haan and Ferreira [22].

4 Simulation study

We examine the behavior of our estimator on several finite-sample situations. To make it easier to showcase our results, we focus on the case $\mathcal{E} = [0,1]$

equipped with the standard absolute value distance. We set, for $x \in \mathcal{E}$, $\gamma(x) = (1 + \sin(2\pi x)/3)/2$. We consider three different models for the conditional distribution function of Y given $X = x$, all of which have conditional tail index $\gamma(x)$. The first one is the Fréchet distribution:

$$F(y|x) = \exp(-y^{-1/\gamma(x)}),$$

for all $y > 0$. For this distribution, $\rho(x) = -1$. The second one is the absolute value of the Student distribution with $1/\gamma(x)$ degrees of freedom: for this distribution, $\rho(x) = -2\gamma(x)$. The third and final one is a Burr distribution, which has distribution function:

$$F(y|x) = 1 - (1 + y^{-\rho(x)/\gamma(x)})^{1/\rho(x)},$$

for all $y > 0$. For this distribution, $\rho(x) = \rho$ is assumed to be constant and we choose $\rho \in \{-3/2, -1, -1/2\}$.

In this simulation study, our goal is to estimate the conditional extreme-value index at the three points $x = 1/4, 1/2$ and $3/4$. The function $\Psi(.|x, u)$ is chosen as $\Psi_\theta(.|u)$, where $\theta \in (0, \infty)$ and:

$$\Psi_\theta(\alpha|u) = \frac{(\theta+1)^2}{\theta u^{\theta+1}} \left(\frac{u^\theta}{\theta+1} - \alpha^\theta \right).$$

In this context, condition **(A2)** is satisfied with

$$\Phi(\alpha|x) =: \Phi_\theta(\alpha) = \frac{\theta+1}{\theta}(1-\alpha^\theta) \quad \text{and} \quad g(u|x) = \frac{\theta+1}{\theta}\left((1-u^\theta)+\theta\right).$$

We choose $\theta = 0.6833$; this value can be seen as a minimizer of (a modified version of) the AMSE of the estimator, see Gardes *et al.* [18].

4.1 A global comparison with other methods

We start by comparing our estimator with the following techniques:

The estimator of Goegebeur *et al.* [20]. This estimator is given by:

$$\widehat{\gamma}^{\mathrm{GGS}}(x, \omega_x, h_x) = \frac{T_n^{(1)}(x, \omega_x, h_x)}{T_n^{(0)}(x, \omega_x, h_x)},$$

where for all $t \geq 0$,

$$T_n^{(t)}(x, \omega_x, h_x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_x} K\left(\frac{x - X_i}{h_x}\right) (\max(0, \log Y_i - \log \omega_x))^t \mathbb{I}\{Y_i > \omega_x\}.$$

In the original paper, the estimator is defined and studied only when the threshold sequence $\omega_x \to \infty$ is *nonrandom*. Thus, we first compute the quantity

$\widehat{\gamma}^{\mathrm{GGS}}(x,\omega_x,h_x)$ with $\omega_x = q(u_x|x)$, $u_x = u_x(n) \to 0$ as $n \to \infty$, but we note that in this case, $\widehat{\gamma}^{\mathrm{GGS}}(x,\omega_x,h_x)$ is not an estimator since ω_x is unknown. As advised in Goegebeur *et al.* [20], we also compare our results with the estimator obtained by setting $\omega_x = \widehat{q}_n(u_x|x,h_x)$, which is actually a *random* threshold sequence. Finally, we let $K(x) = (15/16)(1-x^2)^2\mathbb{I}\{|x| \le 1\}$, corresponding to the biweight kernel.

The generalized Pickands-type estimator of Gardes and Girard [17]. For $J \ge 2$ and $0 < \tau_J < \ldots < \tau_1 < 1$, this estimator is given by:

$$\widehat{\gamma}^{\mathrm{GG}}(x,u_x,h_x) = \sum_{j=1}^{J}(\log \widetilde{q}_n(\tau_j u_x|x,h_x) - \log \widetilde{q}_n(u_x|x,h_x)) \bigg/ \sum_{j=1}^{J}\log(1/\tau_j)$$

where for $u \in (0,1)$, $\widetilde{q}_n(u|x,h_x) = \inf\{y \in \mathbb{R} \mid \widetilde{F}(y|x,h_x) \ge 1-u\}$ with

$$\widetilde{F}(y|x,h_x) = \sum_{i=1}^{n} K\left(\frac{x-X_i}{h_x}\right)\mathbb{I}_{\{Y_i \le y\}} \bigg/ \sum_{i=1}^{n} K\left(\frac{x-X_i}{h_x}\right).$$

Following their advice, we set $J = 10$ and $\tau_j = 1/j^2$. We take K to be the biweight kernel.

We then choose grids of values $\mathcal{H}$ for h_x and $\mathcal{U}$ for $u_x \in (0,1)$. For a given $n-$sample, each estimator is computed for every value of $h_x \in \mathcal{H}$ and $u_x \in \mathcal{U}$ with

$$\mathcal{H} = \{0.05, 0.075, \ldots, 0.35\} \text{ and } \mathcal{U} = \{0.025, 0.05, \ldots, 0.5\}.$$

This procedure is repeated on $S = 1000$ independent replications of an $n-$sample of size $n = 300$ in each of the cases detailed above. Visual comparisons of the mean squared errors (MSEs) of each method at $x = 0.5$ for $(u_x,h_x) \in \mathcal{U} \times \mathcal{H}$ are provided on Figures 1–5.
All in all, it appears that the MSE of our estimator $\widehat{\gamma}$ seems to be fairly stable with respect to (u_x,h_x). In this respect, it appears to perform equally well or better than the other estimators. A second remark is that the MSE of any of the four estimators tends to increase as h_x increases. This was expected since a higher h_x means taking into account observations whose associated covariates are further from x, which can increase the bias of the estimate.

4.2 How to choose u_x and h_x

Of course, in practical situations, a choice of u_x and h_x has to be implemented. With this aim in mind, we introduce the statistic

$$\widehat{C}(x,u_x,h_x) := v_x^{1/2}\frac{\widehat{\gamma}(x,u_x,h_x) - \widehat{\gamma}^H(x,u_x,h_x)}{\widehat{\gamma}(x,u_x,h_x)},$$

where $\widehat{\gamma}^H(x,u_x,h_x)$ is the adaptation of the Hill estimator given in (4). We have the following result:

Proposition 1. *Assume that the hypotheses of Theorem 2 hold. Then, as* $n \to \infty$*:*

$$\widehat{C}(x, u_x, h_x) \xrightarrow{d} \mathcal{N}\left(\int_0^1 (\Phi(\alpha|x) - 1)\alpha^{-\rho(x)} d\alpha, \int_0^1 \Phi^2(\alpha|x) d\alpha - 1 \right).$$

In other words, the relative error $|\widehat{C}(x, u_x, h_x)|$ should not be too large if u_x and h_x are suitably chosen. Motivated by Proposition 1, our procedure is thus the following. For every $u_x \in \mathcal{U}$, we compute the set

$$\mathcal{S}(x, u_x) = \{|\widehat{C}(x, u_x, h_x)|, \; h_x \in \mathcal{H}\}.$$

Let then $s(x, u_x)$ be the median of $\mathcal{S}(x, u_x)$, and compute

$$h_x^*(u_x) = \min\{h_x \in \mathcal{H} \mid |\widehat{C}(x, u_x, h_x)| > s(x, u_x)\}.$$

Next, we compute the set

$$\mathcal{T}(x) = \{|\widehat{C}(x, u_x, h_x^*(u_x))|, \; u_x \in \mathcal{U}\}.$$

Let now $t(x)$ be the median of $\mathcal{T}(x)$, and compute

$$u_x^* = \min\{u_x \in \mathcal{U} \mid |\widehat{C}(x, u_x, h_x^*(u_x))| > t(x)\}.$$

We finally choose $u_x := u_x^*$ and $h_x := h_x^*(u_x^*)$.

Once again, we repeat this procedure on $S = 1000$ independent replications of an $n-$sample of size $n = 300$. Boxplots of the results at each of the three points $x = 0.25, 0.5$ and 0.75 are provided on Figure 6.

The results seem globally satisfying in each case. We remark that for the Burr distribution, the finite sample performance of the method deteriorates as $|\rho(x)|$ decreases. This was expected since $|\rho(x)|$ is the second-order parameter that controls the rate of convergence of the asymptotic bias to 0: the larger is $|\rho(x)|$, the smaller is the order of the asymptotic bias. Moreover, our simulation study shows that in practical situations, our estimator suffers from a finite-sample bias which becomes larger for smaller values of $\rho(x)$. This can be seen as a consequence of Theorem 2, in which it appears that the asymptotic bias of the estimator directly depends on the asymptotic behavior of $\Delta(.|x)$ and thus on its second-order parameter $|\rho(x)|$. We point out that this is actually a common characteristic of many tail index estimators which is due to the extreme-value framework.

5 Real data example

In this section, we study a real hurricane data set. Our data come from the Atlantic Hurricane database (HURDAT2), which is available on the website of the U.S.A. National Weather Service, see `http://www.nhc.noaa.gov/data/`. In particular, we focus on the period starting from January 1st, 1950 to December 31st, 2013. For a given hurricane occurring during this timeframe, we retain the time and location at which the related wind speeds attained their maximum. Our variable of interest is then the maximal wind speed and our covariate is the location. There are 944 observations in our data set, which were recorded in the geographical zone $\mathcal{E} = [98.8°\mathrm{W}, 45°\mathrm{W}] \times [8°\mathrm{N}, 53°\mathrm{N}]$. The set $\mathcal{E}$ is equipped with the classical Euclidean distance.

When dealing with environmental data, one should keep in mind that there are various statistical concerns, such as independence and stationarity. We shall not examine these issues in detail here. We merely point out that retaining the maximal wind speeds, which is standard practice when considering the extremes of univariate random variables, can reasonably be expected to yield independent observations. Furthermore, restricting our study to the timeframe 1950–2013, instead of the period 1851–2013 suggested by the original data set, is in our opinion a step towards ensuring stationarity of the data.

Various studies have considered wind speed data from an extreme value perspective. Among them, we mention Beirlant *et al.* [3] who studied daily maximal wind speed data for three cities in the U.S.A., Brabson and Palutikof [6] who introduced a Generalized Pareto Distribution (GPD) model for extreme wind speeds in Scotland, Coles and Simiu [8] who suggested a GPD model and applied it to a simulated data set for hurricane wind speeds in Miami, Florida, U.S.A., and Jagger and Elsner [25] who took particular climate indicators as covariates in order to study tropical cyclone wind speeds along the U.S.A. coastline. Although the extreme value framework seems to be fairly adapted to the study of extreme wind speeds, there seems to be no general consensus about what type of distribution arises. One the one hand, Coles and Simiu [8] and Jagger and Elsner [25] find that the distributions of wind speeds they study are short-tailed, namely they are bounded from above; on the other hand, Beirlant *et al.* [3] and Brabson and Palutikof [6] find evidence to support that the distribution of wind speeds may be heavy-tailed depending on the location.

Moreover, tail index estimators such as the Hill estimator and their generalizations to the random covariate framework may be used to detect the presence of heavy tails, as shown in de Haan and Ferreira [22, Theorem 3.2.4], as well as lighter tails or even a short-tailed distribution, since it is easy to see that our estimator converges pointwise to 0 provided the conditional distribution has a finite right endpoint and satisfies a continuity property. A conditional tail index estimator such as the one we introduce in this paper can therefore be considered as an exploratory tool to analyze a data set from the extreme value perspective.

We thus compute our estimator, using the selection rule of u_x and h_x detailed in Section 4.2, on a grid of points which are chosen to be sufficiently close to at least one observation in our data set. A qualitative result, superimposed to a map of the North Atlantic region, is given in Figure 7. It can be seen that hurricane wind speeds may indeed be considered heavy-tailed in a large part of the Gulf of Mexico, while they look lighter-tailed elsewhere, for example in the Caribbean Sea. Using light-tailed distributions, for instance one featuring an exponential decay in its right tail, or short-tailed distributions might therefore be more appropriate in the latter region.

6 Concluding remarks

In this paper, we introduced and studied an estimator which is a weighted integral of the standard conditional log-quantile estimator. This class of estimators is fairly flexible; furthermore, particular choices of the weighting function yield generalizations of well-known tail index estimators. The asymptotic properties of our estimator were established and its finite-sample properties were seen to be satisfying.
It was however highlighted that our estimator, as many other tail index estimators do, may suffer from a finite-sample bias which makes it overestimate the conditional tail index. This can be a problem in practice: for example, in actuarial science, overestimating the tail index of the losses means that these losses are thought to have a bigger tail than they have in reality, and thus that they are expected to cost more than they actually should. This, in turn, can force an insurance firm to build bigger reserves than necessary by increasing the premiums of its customers, through which it could lose a portion of the market share. Future research on this topic therefore includes developing a bias-reduced version of our estimator. Moreover, it is often thought that estimating the tail index is the first step before estimating extreme quantiles of a distribution. It would thus be nice to develop a conditional extreme quantile estimator based on our technique and investigate its behavior.

References

[1] Beirlant, J., Dierckx, G., Guillou, A., Stărică, C. (2002) On exponential representations of log-spacings of extreme order statistics, *Extremes* **5**(2): 157–180.

[2] Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J.L. (2004) *Statistics of extremes – Theory and applications*, Wiley Series in Probability and Statistics.

[3] Beirlant, J., Vynckier, P., Teugels, J.L. (1996) Excess functions and estimation of the extreme value index, *Bernoulli* **2**(4): 293–318.

[4] Billingsley, P. (1979) *Probability and measure*, John Wiley and Sons.

[5] Bingham, N.H., Goldie, C.M., Teugels, J.L. (1987) *Regular Variation*, Cambridge, U.K.: Cambridge University Press.

[6] Brabson, B.B., Palutikof, J.P. (2000) Tests of the generalized Pareto distribution for predicting extreme wind speeds, *Journal of Applied Meteorology* **39**: 1627–1640.

[7] Chavez-Demoulin, V., Davison, A.C. (2005) Generalized additive modelling of sample extremes, *Journal of the Royal Statistical Society, series C* **54**: 207–222.

[8] Coles, S., Simiu, E. (2003) Estimating uncertainty in the extreme value analysis of data generated by a hurricane simulation model, *Journal of Engineering Mechanics* **129**(11): 1288–1294.

[9] Daouia, A., Gardes, L., Girard, S., Lekina, A. (2011) Kernel estimators of extreme level curves, *Test* **20**(2): 311–333.

[10] Daouia, A., Gardes, L., Girard, S. (2013) On kernel smoothing for extremal quantile regression, *Bernoulli* **19 5B**: 2557–2589.

[11] Davison, A.C., Ramesh, N.I. (2000) Local likelihood smoothing of sample extremes, *Journal of the Royal Statistical Society, series B* **62**: 191–208.

[12] Davison, A.C., Smith, R.L. (1990) Models for exceedances over high thresholds, *Journal of the Royal Statistical Society, series B* **52**: 393–442.

[13] Dekkers, A.L.M., Einmahl, J.H.J., de Haan, L. (1989) A moment estimator for the index of an extreme-value distribution, *Annals of Statistics* **17**(4): 1833–1855.

[14] Ferraty, F., Vieu, P. (2006) *Nonparametric Functional Data Analysis: Theory and Practice*, Springer Series in Statistics, Springer.

[15] Gardes, L., Girard, S. (2008) A moving window approach for nonparametric estimation of the conditional tail index, *Journal of Multivariate Analysis* **99**: 2368–2388.

[16] Gardes, L., Girard, S. (2010) Conditional extremes from heavy-tailed distributions: an application to the estimation of extreme rainfall return levels, *Extremes* **13**: 177–204.

[17] Gardes, L., Girard, S. (2012) Functional kernel estimators of large conditional quantiles, *Electronic Journal of Statistics* **6**: 1715–1744.

[18] Gardes, L., Guillou, A., Schorgen, A. (2012) Estimating the conditional tail index by integrating a kernel conditional quantile estimator, *Journal of Statistical Planning and Inference* **142**: 1586–1598.

[19] Gardes, L., Stupfler, G. (2014) Estimation of the conditional tail index using a smoothed local Hill estimator, *Extremes* **17**(1): 45–75.

[20] Goegebeur, Y., Guillou, A., Schorgen, A. (2014) Nonparametric regression estimation of conditional tails – the random covariate case, *Statistics* **48**(4): 732–755.

[21] Goegebeur, Y., Guillou, A., Stupfler, G. (2014). Uniform asymptotic properties of a nonparametric regression estimator of conditional tails, *Annales de l'Institut Henri Poincaré (B): Probability and Statistics*, to appear. Available at `http://hal.archives-ouvertes.fr/hal-00794724`.

[22] de Haan, L., Ferreira, A. (2006) *Extreme value theory: An introduction*, Springer.

[23] Hall, P., Tajvidi, N. (2000) Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data, *Statistical Science* **15**: 153–167.

[24] Hill, B.M. (1975) A simple general approach to inference about the tail of a distribution, *Annals of Statistics* **3**: 1163–1174.

[25] Jagger, T.H., Elsner, J.B. (2009) Modeling tropical cyclone intensity with quantile regression, *International Journal of Climatology* **29**: 1351–1361.

[26] Kratz, M., Resnick, S. (1996) The QQ-estimator and heavy tails, *Stochastic Models* **12**: 699–724.

[27] Pickands, J. (1975) Statistical inference using extreme order statistics, *Annals of Statistics* **3**: 119–131.

[28] Resnick, S., Stărică, C. (1997) Smoothing the Hill estimator, *Advances in Applied Probability* **29**: 271–293.

[29] Schultze, J., Steinebach, J. (1996) On least squares estimates of an exponential tail coefficient, *Statistics and Decisions* **14**: 353–372.

[30] Smith, R. L. (1989) Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone (with discussion), *Statistical Science* **4**: 367–393.

[31] Stupfler, G. (2013) A moment estimator for the conditional extreme-value index, *Electronic Journal of Statistics* **7**: 2298–2343.

[32] Wang, H., Tsai, C.L. (2009) Tail index regression, *Journal of the American Statistical Association* **104**(487): 1233–1240.

Appendix

6.1 Auxiliary results and their proofs

The first result is a classical equivalent of $M(x, h_x)$: see also Lemma 1 in Stupfler [31].

Lemma 1. *Pick $x \in \mathbb{R}^d$ and assume that $m_x(h_x) \to \infty$ as $n \to \infty$. Then, for any $\delta > 0$:*

$$[m_x(h_x)]^{(1-\delta)/2} \left| \frac{M(x,h_x)}{m_x(h_x)} - 1 \right| \xrightarrow{\mathbb{P}} 0 \quad as \quad n \to \infty.$$

Proof of Lemma 1. The statement is a straightforward consequence of Chebyshev's inequality. ■

We let $\{Y_i^*, i = 1, \ldots, M(x,h_x)\}$ be the response variables whose associated covariates $\{X_i^*, i = 1, \ldots, M(x,h_x)\}$ belong to the ball $B(x,h_x)$. Lemma 2 below is similar in spirit to Lemma 2 in Stupfler [31] and Lemma 4 in Gardes and Stupfler [19].

Lemma 2. *For any x such that $\mathbb{P}(X \in B(x,h_x)) \neq 0$, given $M(x,h_x) = p \geq 1$, the random variables $V_i = 1 - F(Y_i^*|X_i^*)$, $i = 1, \ldots, p$, are independent standard uniform random variables.*

Proof of Lemma 2. If $(u_1, \ldots, u_p) \in \mathbb{R}^p$, then since the random pairs (X_i, Y_i) are independent and identically distributed, we have:

$$\mathbb{P}\left(\bigcap_{i=1}^{p} \{V_i \leq u_i\}, M(x,h_x) = p \right) = \binom{n}{p} \prod_{i=1}^{p} \varrho(u_i|x,h_x) \prod_{i=p+1}^{n} \mathbb{P}\left(X_i \notin B(x,h_x)\right),$$

where $\varrho(t|x,h_x) := \mathbb{P}(F(Y|X) \geq 1-t, X \in B(x,h_x))$. Furthermore, for all $t \in [0,1]$,

$$\begin{aligned} \varrho(t|x,h_x) &= \int_{B(x,h_x)} \left(\int_{\mathbb{R}} \mathbb{I}\{F(y|x) \geq 1-t\} f(y|x) dy \right) \mathbb{P}_X(dx) \\ &= t\mathbb{P}(X \in B(x,h_x)), \end{aligned}$$

by a change of variables in the inner integral. Since the random variable $M(x,h_x)$ follows a binomial distribution with parameters n and $\mathbb{P}\left(X \in B(x,h_x)\right)$, it follows that:

$$\mathbb{P}\left(\bigcap_{i=1}^{p} \{V_i \leq u_i\} | M(x,h_x) = p \right) = u_1 \ldots u_p,$$

which is the result. ■

Lemma 3 shows that the estimator $\widehat{\gamma}(x,k_x,h_x)$ can be approximated by a weighted Hill estimator (see Hill [24]).

Lemma 3. *Let U_i, $i \geq 1$ be independent standard uniform random variables. For any x such that $\mathbb{P}(X \in B(x,h_x)) \neq 0$, we may write*

$$\widehat{\gamma}(x,u_x,h_x) = \widetilde{\gamma}(x,u_x,h_x) + R(x,u_x,h_x)$$

where the conditional distribution of $\widetilde{\gamma}(x,u_x,h_x)$ given $M(x,h_x)=p$ is that of

$$\overline{\gamma}(x,u_x,p) = \sum_{i=1}^{\lfloor pu_x(p)\rfloor} \left(\frac{1}{i}\int_0^{i/p} \Psi(\alpha|x,u_x(p))d\alpha\right) i \log \frac{q(U_{i,p}|x)}{q(U_{i+1,p}|x)}, \tag{7}$$

and $|R(x,u_x,h_x)| \le \overline{R}(x,u_x,h_x)$ where the conditional distribution of $\overline{R}(x,u_x,h_x)$ given $M(x,h_x)=p$ is that of

$$2\omega(U_{1,p},U_{p,p},x,h_x)\int_0^{u_x(p)} |\Psi(\alpha|x,u_x(p))|d\alpha.$$

Proof of Lemma 3. For the sake of brevity, let us write $M_n := M(x,h_x)$ and $k_x(M_n) = M_n u_x(M_n)$ and let for $p \in \mathbb{N}^*$,

$$w_{i,p}(x) := \int_{(i-1)/p}^{i/p} \Psi(\alpha|x,u_x(p))d\alpha.$$

Since for all $i \in \{1,\ldots,M_n\}$ and $\alpha \in [(i-1)/M_n, i/M_n)$, $\widehat{q}_n(\alpha|x,h_x) = Y^*_{M_n-i+1,M_n}$, we may write:

$$\widehat{\gamma}(x,u_x,h_x) = \sum_{i=1}^{\lfloor k_x(M_n)\rfloor} w_{i,M_n}(x)\log\frac{Y^*_{M_n-i+1,M_n}}{Y^*_{M_n-\lfloor k_x(M_n)\rfloor,M_n}}.$$

Write $\widehat{\gamma}(x,u_x,h_x) = \widetilde{\gamma}(x,u_x,h_x) + R(x,u_x,h_x)$ with $\widetilde{\gamma}(x,u_x,h_x)$ given by:

$$\sum_{i=1}^{\lfloor k_x(M_n)\rfloor} w_{i,M_n}(x)\log\frac{q(1-F(Y^*_{M_n-i+1,M_n}|X^*_{(i)})|x)}{q(1-F(Y^*_{M_n-\lfloor k_x(M_n)\rfloor,M_n}|X^*_{(\lfloor k_x(M_n)\rfloor+1)})|x)},$$

where, for $i=1,\ldots,M_n$, $X^*_{(i)}$ is the covariate associated to $Y^*_{M_n-i+1,M_n}$. Now, given $M_n=p$, Lemma 2 entails that there exist independent standard uniform variables $U_1,\ldots,U_p$ such that the conditional distribution of $\widetilde{\gamma}(x,u_x,h_x)$ given $M_n=p$ is that of

$$\sum_{i=1}^{\lfloor k_x(p)\rfloor} w_{i,p}(x)\log\frac{q(U_{i,p}|x)}{q(U_{\lfloor k_x(M_n)\rfloor+1,p}|x)} = \sum_{i=1}^{\lfloor k_x(p)\rfloor} w_{i,p}(x)\sum_{j=i}^{\lfloor k_x(p)\rfloor}\log\frac{q(U_{j,p}|x)}{q(U_{j+1,p}|x)},$$

which is equal to $\overline{\gamma}(x,u_x,p)$ by switching the summation order. Let us now focus on the term $R(x,u_x,h_x) = \widehat{\gamma}(x,u_x,h_x) - \widetilde{\gamma}(x,u_x,h_x)$. Let $V_i = 1-F(Y_i^*|X_i^*)$. Since $q(.|x)$ is continuous and decreasing, one has, for $i=1,\ldots,M_n$,

$$\begin{aligned}\log q(V_i|x) - \omega(V_{1,M_n},V_{M_n,M_n},x,h_x) &\le \log Y_i^* = \log q(V_i|X_i^*) \\ &\le \log q(V_i|x) + \omega(V_{1,M_n},V_{M_n,M_n},x,h_x).\end{aligned}$$

It follows from Lemma 1 in Gardes and Stupfler [19] that:

$$\left|\log Y^*_{M_n-i+1,M_n} - \log q(V_{i,M_n}|x)\right| \le \omega(V_{1,M_n},V_{M_n,M_n},x,h_x).$$

Hence,

$$\left|\log \frac{Y^*_{M_n-i+1,M_n}}{Y^*_{M_n-\lfloor k_x(M_n)\rfloor,M_n}} - \log \frac{q(V_{i,M_n}|x)}{q(V_{\lfloor k_x(M_n)\rfloor+1,M_n}|x)}\right| \leq 2\omega(V_{1,M_n}, V_{M_n,M_n}, x, h_x),$$

and thus $|R(x, u_x, h_x)|$ is bounded from above by

$$\overline{R}(x, u_x, h_x) := 2\omega(V_{1,M_n}, V_{M_n,M_n}, x, h_x) \int_0^{u_x(M_n)} |\Psi(\alpha|x, u_x(M_n))| d\alpha.$$

Applying Lemma 2 completes the proof. ■

Our next result is dedicated to the study of some particular Riemann sums.

Lemma 4. *Let f be an integrable function on $(0, 1)$. Assume that f is nonnegative and nonincreasing. For any nonnegative continuous function g on $[0, 1]$ and any sequence (m_n) converging to infinity, we have that:*

$$\lim_{n\to\infty} \frac{1}{m_n} \sum_{i=1}^{\lfloor m_n\rfloor} f(i/m_n)g(i/m_n) = \int_0^1 f(t)g(t)dt.$$

If moreover f is square-integrable then:

$$\lim_{n\to\infty} \sqrt{m_n} \left|\frac{1}{m_n} \sum_{i=1}^{\lfloor m_n\rfloor} f(i/m_n) - \int_0^1 f(t)dt\right| = 0.$$

Proof of Lemma 4. Define

$$S_n(f, g) := \frac{1}{m_n} \sum_{i=1}^{\lfloor m_n\rfloor} f(i/m_n)g(i/m_n) \text{ and } S(f, g) := \int_0^1 f(t)g(t)dt.$$

Note first that:

$$|S(f, g) - S_n(f, g)| \leq \sum_{i=1}^{\lfloor m_n\rfloor} \int_{(i-1)/\lfloor m_n\rfloor}^{i/\lfloor m_n\rfloor} \left|f(t)g(t) - f(i/m_n)g(i/m_n)\frac{\lfloor m_n\rfloor}{m_n}\right| dt.$$

Since g is nonnegative on $[0, 1]$ and f is nonincreasing, it is straightforward that for all $t \in [(i - 1)/\lfloor m_n\rfloor, i/\lfloor m_n\rfloor)$

$$\begin{aligned}\left|f(t)g(t) - f(i/m_n)g(i/m_n)\frac{\lfloor m_n\rfloor}{m_n}\right| &\leq f(t) \sup_{|s-s'|\leq 1/m_n} |g(s) - g(s')| \\ &+ \|g\|_\infty f(t)\left(1 - \frac{\lfloor m_n\rfloor}{m_n}\right) \\ &+ \|g\|_\infty (f(t) - f(i/m_n))\end{aligned}$$

where $\|g\|_\infty$ is the finite supremum of g on $[0,1]$. Using the fact that, since f is nonincreasing, one has for $i=2,\ldots,\lfloor m_n\rfloor$ that $f(t)-f(i/m_n)\le f((i-1)/m_n)-f(i/m_n)$, the previous inequality leads to

$$\begin{aligned}|S(f,g)-S_n(f,g)| &\le \int_0^1 f(t)dt \sup_{|s-s'|\le 1/n}|g(s)-g(s')| \\ &+ \|g\|_\infty \int_0^1 f(t)dt\left(1-\frac{\lfloor m_n\rfloor}{m_n}\right) \\ &+ \|g\|_\infty\left(\int_0^{1/\lfloor m_n\rfloor} f(t)dt-\frac{f(1)}{\lfloor m_n\rfloor}\right)\to 0 \qquad (8)\end{aligned}$$

by the uniform continuity of g on $[0,1]$. This proves the first statement of the result. To prove the second one, take $g=1$ in (8) to get:

$$\begin{aligned}(m_n)^{1/2}|S(f,1)-S_n(f,1)| &\le (m_n)^{1/2}\left(1-\frac{\lfloor m_n\rfloor}{m_n}\right)\int_0^1 f(t)dt \\ &+ (m_n)^{1/2}\int_0^{1/\lfloor m_n\rfloor} f(t)dt.\end{aligned}$$

Since $1-\lfloor m_n\rfloor/m_n<1/m_n$, the first term of the right-hand side converges to 0. By the Cauchy-Schwarz inequality,

$$(m_n)^{1/2}\int_0^{1/\lfloor m_n\rfloor} f(t)dt\le\left(\frac{m_n}{\lfloor m_n\rfloor}\right)^{1/2}\left(\int_0^{1/\lfloor m_n\rfloor} f^2(t)dt\right)^{1/2}\to 0,$$

since f^2 is integrable on $(0,1)$. The proof is complete. ■

Lemma 5 examines the asymptotic properties (as $p\to\infty$) of the quantity $\overline{\gamma}(x,u_x,p)$ introduced in Lemma 3, equation (7):

$$\overline{\gamma}(x,u_x,p)=\sum_{i=1}^{\lfloor pu_x(p)\rfloor}\left(\frac{1}{i}\int_0^{i/p}\Psi(\alpha|x,u_x(p))d\alpha\right) i\log\frac{q(U_{i,p}|x)}{q(U_{i+1,p}|x)},$$

where $U_1,\ldots,U_p$ are independent standard uniform random variables. Recall from Theorem 2 the notations

$$\mathcal{AB}_x(\Phi,\rho(x))=\int_0^1\Phi(\alpha|x)\alpha^{-\rho(x)}d\alpha \quad\text{and}\quad \mathcal{AV}_x(\Phi)=\int_0^1\Phi^2(\alpha|x)d\alpha.$$

Lemma 5. *Assume that $u_x\in\mathcal{RV}_{-a(x)}$ with $a(x)\in(0,1)$.*

i) If **(M1)** *and* **(A1)** *hold, then $\overline{\gamma}(x,u_x,p)\xrightarrow{\mathbb{P}}\gamma(x)$.*

ii) If **(M2)** *and* **(A2)** *hold and $(zu_x(z))^{1/2}\Delta(1/u_x(z)|x)\to\lambda(x)\in\mathbb{R}$ as z goes to infinity then:*

$$(pu_x(p))^{1/2}(\overline{\gamma}(x,u_x,p)-\gamma(x))\xrightarrow{d}\mathcal{N}\left(\lambda(x)\mathcal{AB}_x(\Phi,\rho(x)),\gamma^2(x)\mathcal{AV}_x(\Phi)\right).$$

Proof of Lemma 5. Pick $p \geq 2$ and let for $i \in \{1, \ldots, \lfloor pu_x(p) \rfloor\}$:

$$w_{i,p}(x) = \frac{1}{i} \int_0^{i/p} \Psi(\alpha | x, u_x(p)) d\alpha.$$

i) To show the consistency statement, we set $E_i(p) = i \log(U_{i+1,p}/U_{i,p})$ and we use model **(M1)** to rewrite $\overline{\gamma}(x, u_x, p) - \gamma(x)$ as:

$$\overline{\gamma}(x, u_x, p) - \gamma(x) = S_{1,p}(x) + S_{2,p}(x) + S_{3,p}(x) + S_{4,p}(x), \tag{9}$$

with

$$\begin{aligned}
S_{1,p}(x) &= \gamma(x) \sum_{i=1}^{\lfloor pu_x(p) \rfloor} w_{i,p}(x)(E_i(p) - 1), \\
S_{2,p}(x) &= \gamma(x) \left(\sum_{i=1}^{\lfloor pu_x(p) \rfloor} w_{i,p}(x) \right) - \gamma(x), \\
S_{3,p}(x) &= \sum_{i=1}^{\lfloor pu_x(p) \rfloor} w_{i,p}(x) \left(i \log \frac{c(U_{i,p}^{-1}|x)}{c(U_{i+1,p}^{-1}|x)} \right) \\
\text{and } S_{4,p}(x) &= \sum_{i=1}^{\lfloor pu_x(p) \rfloor} w_{i,p}(x) \left(i \int_{U_{i+1,p}^{-1}}^{U_{i,p}^{-1}} \frac{\Delta(v|x)}{v} dv \right).
\end{aligned}$$

It is thus enough to show that for any $j \in \{1, 2, 3, 4\}$, $S_{j,p}(x) \xrightarrow{\mathbb{P}} 0$ as $p \to \infty$. We start by controlling the sum $S_{1,p}(x)$: since the random variables $-\log U_i$ are independent standard exponential random variables, Rényi's representation (see de Haan and Ferreira [22]) entails that the $E_i(p)$ are independent standard exponential random variables as well. Thus $S_{1,p}(x)$ is centered and

$$\operatorname{Var}(S_{1,p}(x)) = \gamma^2(x) \sum_{i=1}^{\lfloor pu_x(p) \rfloor} w_{i,p}^2(x). \tag{10}$$

Condition **(A1)** yields:

$$w_{i,p}(x) = \frac{1}{i} \int_0^{i/p} \Psi(\alpha | x, u_x(p)) d\alpha = \frac{1}{pu_x(p)} \Phi\left(i/(pu_x(p)) | x\right). \tag{11}$$

Thus, for any $a \geq 1$ such that $\Phi^a(.|x)$ is integrable on $(0, 1)$:

$$\begin{aligned}
\sum_{i=1}^{\lfloor pu_x(p) \rfloor} w_{i,p}^a(x) &= (pu_x(p))^{1-a} \left(\frac{1}{pu_x(p)} \sum_{i=1}^{\lfloor pu_x(p) \rfloor} \Phi^a\left(i/(pu_x(p)) | x\right) \right) \\
&= (pu_x(p))^{1-a} \int_0^1 \Phi^a(\alpha | x) d\alpha (1 + \mathrm{o}(1)),
\end{aligned} \tag{12}$$

using Lemma 4 with $f = \Phi^a(.|x)$ and $g = 1$. Apply (10) together with (12) for $a = 2$ to get as $p \to \infty$:

$$S_{1,p}(x) \xrightarrow{\mathbb{P}} 0. \tag{13}$$

The nonrandom term $S_{2,p}(x)$ is controlled by using (12) with $a = 1$:

$$S_{2,p}(x) = \gamma(x) \left(\sum_{i=1}^{\lfloor pu_x(p) \rfloor} w_{i,p}(x) - \int_0^1 \Phi(\alpha|x) d\alpha \right) \to 0, \tag{14}$$

as $p \to \infty$. The sum $S_{3,p}(x)$ is controlled by rewriting it as:

$$S_{3,p}(x) = \sum_{j=1}^{\lfloor pu_x(p) \rfloor} \left(\int_{(j-1)/p}^{j/p} \Psi(\alpha|x, u_x(p)) d\alpha \right) \log \frac{c(U_{j,p}^{-1}|x)}{c(U_{\lfloor pu_x(p) \rfloor+1,p}^{-1}|x)}.$$

From this, we deduce that:

$$|S_{3,p}(x)| \le \int_0^{u_x(p)} |\Psi(\alpha|x, u_x(p))| d\alpha \sup_{s,t \ge U_{\lfloor pu_x(p) \rfloor+1,p}^{-1}} \left| \log \frac{c(s|x)}{c(t|x)} \right|,$$

which we use together with condition **(A1)**, the convergence of $c(.|x)$ to a positive constant and the convergence $[u_x(p)]^{-1} U_{\lfloor pu_x(p) \rfloor+1,p} \xrightarrow{\mathbb{P}} 1$ as $p \to \infty$ to get:

$$S_{3,p}(x) \xrightarrow{\mathbb{P}} 0. \tag{15}$$

Finally, to control $S_{4,p}(x)$ we write:

$$|S_{4,p}(x)| \le \left(1 + \frac{S_{1,p}(x) + S_{2,p}(x)}{\gamma(x)} \right) \sup_{v \ge U_{\lfloor pu_x(p) \rfloor+1,p}^{-1}} |\Delta(v|x)|. \tag{16}$$

Use (16) together with (13), (14), the convergence $[u_x(p)]^{-1} U_{\lfloor pu_x(p) \rfloor+1,p} \xrightarrow{\mathbb{P}} 1$ as $p \to \infty$ and the convergence of $|\Delta(.|x)|$ to zero to obtain:

$$S_{4,p}(x) \xrightarrow{\mathbb{P}} 0. \tag{17}$$

Combining (13), (14), (15) and (17) completes the proof of the consistency statement.

ii) To prove the asymptotic normality statement, we note that since **(M2)** holds, we may apply Theorem 2.1 in Beirlant *et al.* [1] to obtain that the random vector $\{i \log(q(U_{i,p}|x)/q(U_{i+1,p}|x)),\ i \in H_{p,x}\}$ where $H_{p,x} := \{1, \ldots, \lfloor pu_x(p) \rfloor\}$ has the same distribution as:

$$\left\{ \left[\gamma(x) + \Delta_{p,x} \left(\frac{i}{\lfloor pu_x(p) \rfloor + 1} \right)^{-\rho(x)} \right] E_i(p) + \nu_{i,p}(x) + \mathrm{o}_{\mathbb{P}}(\Delta_{p,x}),\ i \in H_{p,x} \right\},$$

with $\Delta_{p,x} := \Delta(p/\lfloor pu_x(p)\rfloor | x)$ and where the $\nu_{i,p}(x)$ satisfy

$$\sum_{j=i}^{\lfloor pu_x(p)\rfloor} \frac{|\nu_{j,p}(x)|}{j} = \mathrm{o}_{\mathbb{P}}\left(|\Delta_{p,x}| \max\left(\log\frac{\lfloor pu_x(p)\rfloor+1}{i}, 1\right)\right), \tag{18}$$

uniformly in $i \in H_{p,x}$. Using the definitions of $S_{1,p}(x)$ and $S_{2,p}(x)$ introduced above, we may therefore write:

$$\begin{aligned} & (pu_x(p))^{1/2}\left(\overline{\gamma}(x,u_x,p) - \gamma(x) - \Delta_{p,x}\int_0^1 \Phi(\alpha|x)\alpha^{-\rho(x)}d\alpha\right) \\ \stackrel{d}{=}\; & (pu_x(p))^{1/2}(S_{1,p}(x)+S_{2,p}(x)+S'_{1,p}(x)+S'_{2,p}(x)+S'_{3,p}(x)) + \mathrm{o}_{\mathbb{P}}(1) \end{aligned}$$

with

$$\begin{aligned} S'_{1,p}(x) &= \Delta_{p,x}\sum_{i=1}^{\lfloor pu_x(p)\rfloor} w_{i,p}(x)\left(\frac{i}{\lfloor pu_x(p)\rfloor+1}\right)^{-\rho(x)}(E_i(p)-1), \\ S'_{2,p}(x) &= \Delta_{p,x}\left(\sum_{i=1}^{\lfloor pu_x(p)\rfloor} w_{i,p}(x)\left(\frac{i}{\lfloor pu_x(p)\rfloor+1}\right)^{-\rho(x)} - \int_0^1 \Phi(\alpha|x)\alpha^{-\rho(x)}d\alpha\right), \\ S'_{3,p}(x) &= \sum_{i=1}^{\lfloor pu_x(p)\rfloor} w_{i,p}(x)\nu_{i,p}(x). \end{aligned}$$

We start by examining the convergence of $(pu_x(p))^{1/2}S_{1,p}(x)$. Define $T_{i,p}(x) = w_{i,p}(x)(E_i(p)-1)$ and remark that the $T_{i,p}(x)$, $i \in H_{p,x}$ are independent centered random variables such that

$$S_{1,p}(x) = \gamma(x)\sum_{i=1}^{\lfloor pu_x(p)\rfloor} T_{i,p}(x).$$

By (12) with $a=2$:

$$\mathrm{Var}(S_{1,p}(x)) = \gamma^2(x)\sum_{i=1}^{\lfloor pu_x(p)\rfloor} w_{i,p}^2(x) = [pu_x(p)]^{-1}\gamma^2(x)\int_0^1 \Phi^2(\alpha|x)d\alpha(1+\mathrm{o}(1)),$$

and, by (12) with $a = 2+\kappa$:

$$\begin{aligned} \sum_{i=1}^{\lfloor pu_x(p)\rfloor} \mathbb{E}(|T_{i,p}(x)|^{2+\kappa}) &= \gamma^{2+\kappa}(x)\sum_{i=1}^{\lfloor pu_x(p)\rfloor} w_{i,p}^{2+\kappa}(x)\mathbb{E}(|E_i(p)-1|^{2+\kappa}) \\ &= \mathrm{O}\left([pu_x(p)]^{-1-\kappa}\right). \end{aligned}$$

As a consequence:

$$\frac{1}{[\mathrm{Var}(S_{1,p}(x))]^{1+\kappa/2}}\sum_{i=1}^{\lfloor pu_x(p)\rfloor} \mathbb{E}(|T_{i,p}(x)|^{2+\kappa}) = \mathrm{O}([pu_x(p)]^{-\kappa/2}) \to 0,$$

as $p \to \infty$. Lyapunov's central limit theorem (see Billingsley [4]) thus entails

$$(pu_x(p))^{1/2} S_{1,p}(x) \xrightarrow{d} \mathcal{N}\left(0, \gamma^2(x) \int_0^1 \Phi^2(\alpha|x) d\alpha\right). \tag{19}$$

To control $(pu_x(p))^{1/2} S_{2,p}(x)$, use the second statement of Lemma 4 with $f = \Phi(.|x)$ to obtain:

$$(pu_x(p))^{1/2} S_{2,p}(x) \xrightarrow{\mathbb{P}} 0. \tag{20}$$

To control $(pu_x(p))^{1/2} S'_{1,p}(x)$ we note that since

$$(pu_x(p))^{1/2} \Delta_{p,x} = (pu_x(p))^{1/2} \Delta(1/u_x(p)|x)(1 + \mathrm{o}(1)) \to \lambda(x),$$

we have $|S'_{1,p}(x)| = \mathrm{O}_{\mathbb{P}}(|S''_{1,p}(x)|)$, with

$$S''_{1,p}(x) = (pu_x(p))^{-1/2} \sum_{i=1}^{\lfloor pu_x(p) \rfloor} w_{i,p}(x) \left(\frac{i}{pu_x(p)}\right)^{-\rho(x)} (E_i(p) - 1).$$

The variance of the centered sum $S''_{1,p}(x)$ is such that:

$$\begin{aligned} pu_x(p) \operatorname{Var}(S''_{1,p}(x)) &= \sum_{i=1}^{\lfloor pu_x(p) \rfloor} w_{i,p}^2(x) \left(\frac{i}{pu_x(p)}\right)^{-2\rho(x)} \\ &= (pu_x(p))^{-1} \int_0^1 \Phi(\alpha|x) \alpha^{-2\rho(x)} d\alpha (1 + \mathrm{o}(1)), \end{aligned}$$

where (11) and Lemma 4 were used, with $f = \Phi(.|x)$ and $g : t \mapsto t^{-2\rho(x)}$. As a consequence:

$$(pu_x(p))^{1/2} S'_{1,p}(x) \xrightarrow{\mathbb{P}} 0. \tag{21}$$

The term $(pu_x(p))^{1/2} S'_{2,p}(x)$ is controlled in the following way: we note that $|S'_{2,p}(x)| = \mathrm{O}(|S''_{2,p}(x)|)$ with

$$S''_{2,p}(x) = (pu_x(p))^{-1/2} \left| \sum_{i=1}^{\lfloor pu_x(p) \rfloor} w_{i,p}(x) \left(\frac{i}{\lfloor pu_x(p) \rfloor + 1}\right)^{-\rho(x)} - \int_0^1 \Phi(\alpha|x) \alpha^{-\rho(x)} d\alpha \right|$$

and we use once again (11) and Lemma 4 with $f = \Phi(.|x)$ and $g : t \mapsto t^{-\rho(x)}$ to get that $(pu_x(p))^{1/2} S''_{2,p}(x) \to 0$. Thus:

$$(pu_x(p))^{1/2} S'_{2,p}(x) \to 0. \tag{22}$$

Finally, we use (11) to bound $|S'_{3,p}(x)|$ by:

$$(pu_x(p))^{-1} \sum_{i=1}^{\lfloor pu_x(p) \rfloor} \left| i\Phi\left(\frac{i}{pu_x(p)}|x\right) - (i-1)\Phi\left(\frac{i-1}{pu_x(p)}|x\right) \right| \sum_{j=i}^{\lfloor pu_x(p) \rfloor} \frac{|\nu_{j,p}(x)|}{j}.$$

Using conditions **(A2)** and (18), we thus get:

$$\begin{aligned}|S'_{3,p}(x)| &= \mathrm{o}_{\mathbb{P}}\left(|\Delta_{p,x}|(pu_x(p))^{-1}\sum_{i=1}^{\lfloor pu_x(p)\rfloor} g\left(i/(pu_x(p))|x\right)\max\left(\log\frac{pu_x(p)}{i},1\right)\right)\\ &= \mathrm{o}_{\mathbb{P}}(|\Delta_{p,x}|),\end{aligned}$$

by Lemma 4 with $f=\max(\log(1/.),1)$ and $g=g(.|x)$ if $g(\cdot|x)$ is continuous on $[0,1]$, or $f=\max(\log(1/.),1)g(.|x)$ and $g=1$ if $g(\cdot|x)$ is nonincreasing on $(0,1)$. Consequently:

$$(pu_x(p))^{1/2}S'_{3,p}(x)\xrightarrow{\mathbb{P}}0 \text{ as } p\to\infty. \tag{23}$$

Combining (19), (20), (21), (22) and (23) completes the proof. ■

The ultimate result is a general "de-conditioning" result which is the cornerstone to prove Theorems 1 and 2.

Lemma 6. *Let $(N=N_n)$ be a nonnegative sequence of integer-valued random variables and $(\widetilde{Z}_n)$, (R_n) be two sequences of real-valued random variables. Assume that there exists a sequence of random variables $(\overline{Z}(p))$ such that for any $p\in\mathbb{N}\setminus\{0\}$, the distribution of $\widetilde{Z}_n$ given $N=p$ is that of $\overline{Z}(p)$. Assume also that there exist a nonrandom positive sequence (p_n) of integers tending to infinity and a nonrandom positive sequence (ε_n) converging to 0 such that if $I_n=[p_n(1-\varepsilon_n),p_n(1+\varepsilon_n)]$, we have that $\mathbb{P}(N\notin I_n)\to 0$ as $n\to\infty$. Let $Z_n:=\widetilde{Z}_n+R_n$.*

i) If $\overline{Z}(p)$ converges in probability to 0 as $p\to\infty$ and if for all $t>0$,

$$\lim_{n\to\infty}\sup_{p\in I_n}\mathbb{P}(|R_n|>t|N=p)=0,$$

then (Z_n) converges in probability to 0.

ii) If there exists a positive function $v(.)$ for which $v(p_n)\to\infty$ and $(v(p_n)\overline{Z}(p_n))$ converges in distribution to some absolutely continuous distribution H as $n\to\infty$ and such that

$$\sup_{p,p'\in I_n}\left|\frac{v(p)}{v(p')}-1\right|\to 0 \quad \textit{and} \quad \lim_{n\to\infty}\sup_{p\in I_n}\mathbb{P}(v(p)|R_n|>t|N=p)=0,$$

for all $t>0$ then $(v(p_n)Z_n)$ converges in distribution to H.

Proof of Lemma 6. *i)* To prove the consistency statement, pick $t>0$ and write:

$$\begin{aligned}\mathbb{P}(|Z_n|>t) &= \sum_{j=0}^{\infty}\mathbb{P}(|Z_n|>t|N=j)\,\mathbb{P}(N=j)\\ &\le \sup_{p\in I_n}\mathbb{P}(|Z_n|>t|N=p)+\mathrm{o}(1)\\ &\le \sup_{p\in I_n}\mathbb{P}\left(|\widetilde{Z}_n|>t/2|N=p\right)+\sup_{p\in I_n}\mathbb{P}(|R_n|>t/2|N=p)+\mathrm{o}(1)\end{aligned}$$

as $n \to \infty$. The result follows by noting that:

$$\lim_{n\to\infty} \sup_{p\in I_n} \mathbb{P}(|R_n| > t/2|N = p) = 0,$$

and

$$\lim_{n\to\infty} \sup_{p\in I_n} \mathbb{P}(|\widetilde{Z}_n| > t/2|N = p) = \lim_{n\to\infty} \sup_{p\in I_n} \mathbb{P}(|\overline{Z}(p)| > t/2) = 0.$$

ii) Use first the condition on $v(.)$ to obtain $v(N) = v(p_n)(1 + o_{\mathbb{P}}(1))$. It is therefore enough to prove that $(v(N)Z_n)$ converges in distribution to H. We have for any $t \in \mathbb{R}$ and any $\varepsilon > 0$:

$$\begin{aligned}
|\mathbb{P}(v(N)Z_n \le t) - H(t)| &\le \sum_{j=0}^{\infty} |\mathbb{P}(v(j)Z_n \le t|N = j) - H(t)|\,\mathbb{P}(N = j) \\
&\le \sup_{p\in I_n} |\mathbb{P}(v(p)Z_n \le t|N = p) - H(t)| + \varepsilon/4,
\end{aligned}$$

for n large enough. Since H is continuous, one can find $\kappa > 0$ such that $H(t + \kappa) - H(t - \kappa) \le \varepsilon/8$. Observe that $\sup_{p\in I_n} |\mathbb{P}(v(p)Z_n \le t|N = p) - H(t)| \le T_{1,n} + T_{2,n}$ where:

$$\begin{aligned}
T_{1,n} &= \sup_{p\in I_n} \left|\mathbb{P}(v(p)\widetilde{Z}_n \le t - v(p)R_n, v(p)|R_n| \le \kappa|N = p) - H(t)\right|, \\
T_{2,n} &= \sup_{p\in I_n} \mathbb{P}(v(p)\widetilde{Z}_n \le t - v(p)R_n, v(p)|R_n| > \kappa|N = p).
\end{aligned}$$

By assumption, for n large enough, $T_{2,n} \le \varepsilon/4$ and

$$\begin{aligned}
T_{1,n} &\le \sup_{p\in I_n} \left|\mathbb{P}(v(p)\widetilde{Z}_n \le t + \kappa|N = p) - H(t + \kappa)\right| + (H(t + \kappa) - H(t)) \\
&+ \sup_{p\in I_n} \left|\mathbb{P}(v(p)\widetilde{Z}_n \le t - \kappa|N = p) - H(t - \kappa)\right| + (H(t) - H(t - \kappa)) \\
&\le \sup_{p\in I_n} \left|\mathbb{P}(v(p)\overline{Z}(p) \le t + \kappa) - H(t + \kappa)\right| \\
&+ \sup_{p\in I_n} \left|\mathbb{P}(v(p)\overline{Z}(p) \le t - \kappa) - H(t - \kappa)\right| + \varepsilon/4 \\
&\le \varepsilon/2,
\end{aligned}$$

since $(v(p)\overline{Z}(p))$ converges in distribution to H. Hence

$$\sup_{p\in I_n} |\mathbb{P}(v(p)Z_n \le t|N = p) - H(t)| \le 3\varepsilon/4,$$

which concludes the proof. ■

6.2 Proofs of the main results

Proof of Theorem 1. The main idea is to apply Lemma 6, with $N = M(x, h_x)$, $p_n = \lfloor m_x(h_x) \rfloor$, $\varepsilon_n = p_n^{-1/4}$, $Z_n = \widehat{\gamma}(x, u_x, h_x) - \gamma(x)$, $\widetilde{Z}_n = \widetilde{\gamma}(x, u_x, h_x) - \gamma(x)$, $R_n = \widehat{\gamma}(x, u_x, h_x) - \widetilde{\gamma}(x, u_x, h_x)$ and $\overline{Z}(p) = \overline{\gamma}(x, u_x, p) - \gamma(x)$ with the notation of Lemmas 3 and 5. We observe that Lemma 1 entails $\mathbb{P}(N \notin I_n) \to 0$ as $n \to \infty$. Moreover, from condition **(A1)**, we have that

$$C := \limsup_{u \downarrow 0} \int_0^u |\Psi(\alpha|x, u)| d\alpha < \infty.$$

Apply then Lemma 3 to get for any $t > 0$:

$$\sup_{p \in I_n} \mathbb{P}(|R_n| > t | N = p) \leq \sup_{p \in I_n} \mathbb{P}\left(\omega(U_{1,p}, U_{p,p}, x, h_x) > t/4C\right), \tag{24}$$

where $U_1, \ldots, U_p$ are independent standard uniform random variables. For n large enough, condition (5) thus yields:

$$\begin{aligned} \mathbb{P}(\omega(U_{1,p}, U_{p,p}, x, h_x) > t/4C) &\leq \mathbb{P}(U_{1,p} < [m_x(h_x)]^{-1-\delta}) \\ &+ \mathbb{P}(U_{p,p} > 1 - [m_x(h_x)]^{-1-\delta}). \end{aligned}$$

Since for n large enough:

$$\begin{aligned} &\sup_{p \in I_n} \left[\mathbb{P}(U_{1,p} < [m_x(h_x)]^{-1-\delta}) + \mathbb{P}(U_{p,p} > 1 - [m_x(h_x)]^{-1-\delta})\right] \\ = \ & 2 \sup_{p \in I_n} \left[1 - [1 - [m_x(h_x)]^{-1-\delta}]^p\right] \\ \leq \ & 2\left(1 - [1 - [m_x(h_x)]^{-1-\delta}]^{2m_x(h_x)}\right) \to 0 \end{aligned} \tag{25}$$

as $n \to \infty$, we obtain $\mathbb{P}(|R_n| > t | N = p) \to 0$ for any $t > 0$, uniformly in $p \in I_n$. Finally, the convergence in probability of $(\overline{Z}(p_n))$ to 0 is a consequence of Lemma 5. Applying Lemma 6 completes the proof. ∎

Proof of Theorem 2. Our aim is to apply Lemma 6, with $N = M(x, h_x)$, $p_n = \lfloor m_x(h_x) \rfloor$, $\varepsilon_n = p_n^{-1/4}$, $Z_n = \widehat{\gamma}(x, u_x, h_x) - \gamma(x)$, $\widetilde{Z}_n = \widetilde{\gamma}(x, u_x, h_x) - \gamma(x)$, $R_n = \widehat{\gamma}(x, u_x, h_x) - \widetilde{\gamma}(x, u_x, h_x)$, $\overline{Z}(p) = \overline{\gamma}(x, u_x, p) - \gamma(x)$ with the notation of Lemmas 3 and 5, and $v(p) = (pu_x(p))^{1/2}$. We first observe that Lemma 1 yields $\mathbb{P}(N \notin I_n) \to 0$ as $n \to \infty$. Next, Lemma 3 and condition (6) yield for any $t > 0$ and n large enough:

$$\begin{aligned} &\sup_{p \in I_n} \mathbb{P}(v(p)|R_n| > t | N = p) \\ \leq \ & \sup_{p \in I_n} \mathbb{P}(v(m_x(h_x))\omega(U_{1,p}, U_{p,p}, x, h_x) > t/8C) \\ \leq \ & \sup_{p \in I_n} \left[\mathbb{P}(U_{1,p} < [m_x(h_x)]^{-1-\delta}) + \mathbb{P}(U_{p,p} > 1 - [m_x(h_x)]^{-1-\delta})\right]. \end{aligned}$$

It is then a consequence of (25) that the right-hand side above converges to 0 as $n \to \infty$. Finally, by Lemma 5, the sequence $(v(p_n)\overline{Z}(p_n))$ converges in distribution to the Gaussian distribution with mean $\lambda(x)\mathcal{AB}_x(\Phi, \rho(x))$ and variance $\gamma^2(x)\mathcal{AV}_x(\Phi)$. Applying Lemma 6 completes the proof. ∎

Proof of Proposition 1. We follow the lines of the proof of Theorem 2 and use the Cramér-Wold device to get:

$$v_x^{1/2}\begin{pmatrix} \widehat{\gamma}^H(x,u_x,h_x)-\gamma(x) \\ \widehat{\gamma}(x,u_x,h_x)-\gamma(x) \end{pmatrix} \xrightarrow{d} \mathcal{N}(\mu_x(\Phi,\rho(x)),\Sigma_x(\Phi)),$$

with

$$\mu_x(\Phi,\rho(x)) := \begin{pmatrix} (1-\rho(x))^{-1} \\ \int_0^1 \Phi(\alpha|x)\alpha^{-\rho(x)}d\alpha \end{pmatrix} \text{ and } \Sigma_x(\Phi) := \begin{pmatrix} 1 & 1 \\ 1 & \int_0^1 \Phi^2(\alpha|x)d\alpha \end{pmatrix}.$$

The result is then a consequence of the delta-method. ■

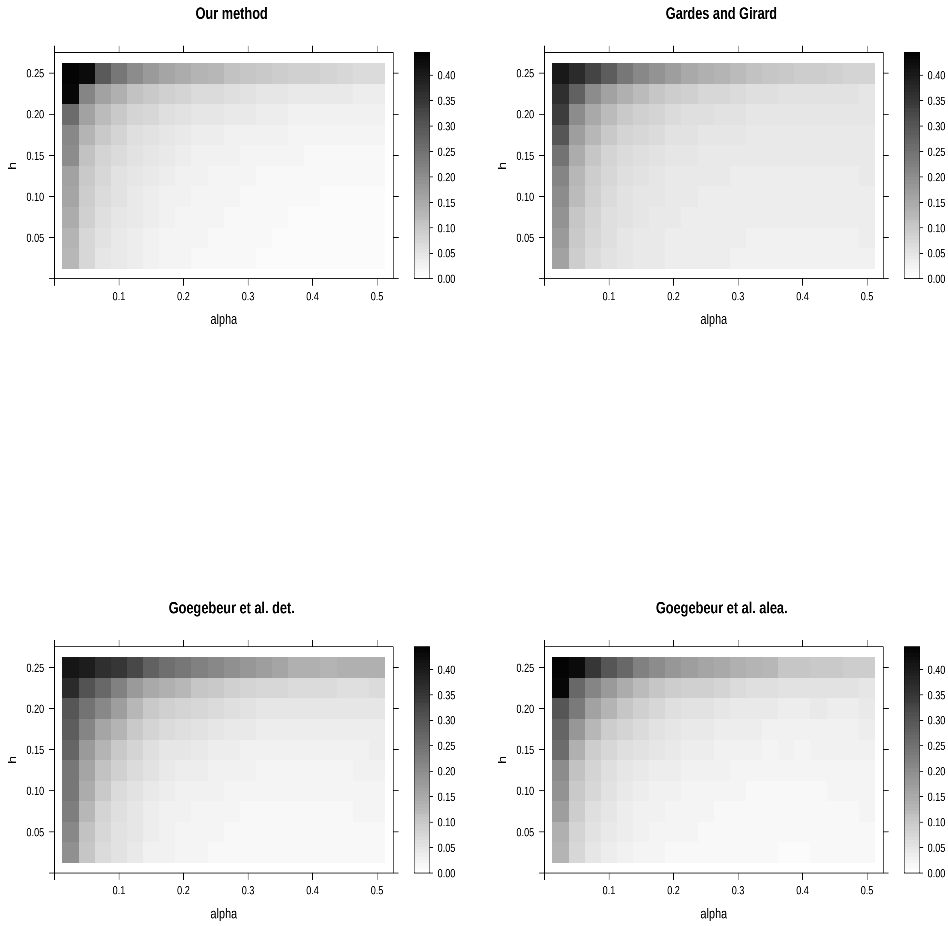

Figure 1: Comparison of the MSE as a function of u_x and h_x for the Burr distribution with $\rho = -3/2$ at $x = 0.5$. Top left: our estimator $\widehat{\gamma}$, top right: estimator $\widehat{\gamma}^{\mathrm{GG}}$ of Gardes and Girard [17], bottom left: estimator $\widehat{\gamma}^{\mathrm{GGS}}$ of Goegebeur *et al.* [20] with $\omega_x = q(u_x|x)$, bottom right: estimator $\widehat{\gamma}^{\mathrm{GGS}}$ of Goegebeur *et al.* [20] with $\omega_x = \widehat{q}(u_x|x, h_x)$.

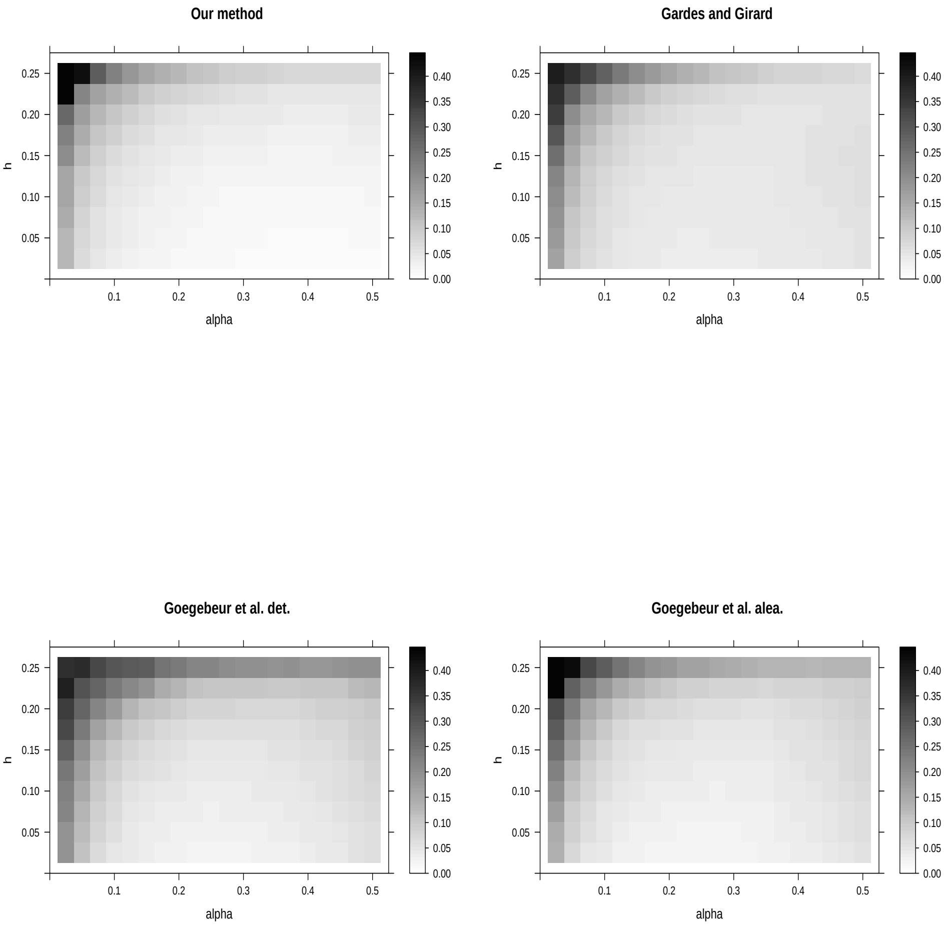

Figure 2: Comparison of the MSE as a function of u_x and h_x for the Burr distribution with $\rho = -1$ at $x = 0.5$. Top left: our estimator $\widehat{\gamma}$, top right: estimator $\widehat{\gamma}^{\mathrm{GG}}$ of Gardes and Girard [17], bottom left: estimator $\widehat{\gamma}^{\mathrm{GGS}}$ of Goegebeur *et al.* [20] with $\omega_x = q(u_x|x)$, bottom right: estimator $\widehat{\gamma}^{\mathrm{GGS}}$ of Goegebeur *et al.* [20] with $\omega_x = \widehat{q}(u_x|x, h_x)$.

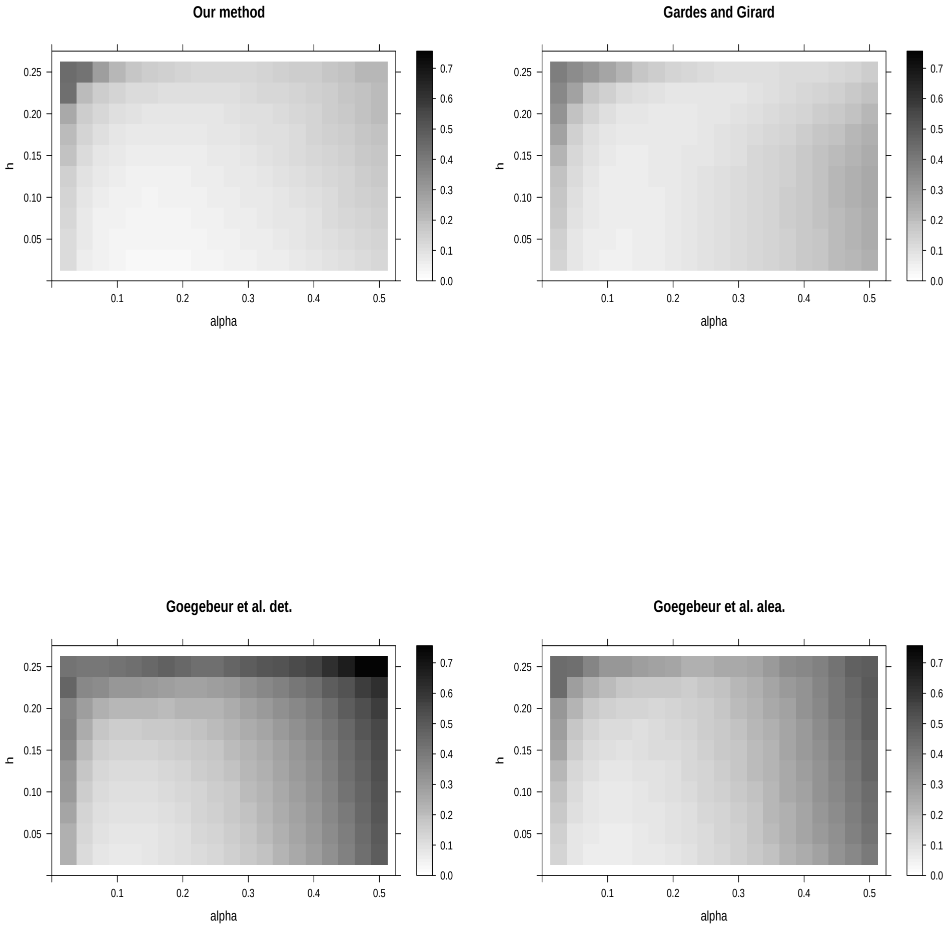

Figure 3: Comparison of the MSE as a function of u_x and h_x for the Burr distribution with $\rho = -1/2$ at $x = 0.5$. Top left: our estimator $\widehat{\gamma}$, top right: estimator $\widehat{\gamma}^{\mathrm{GG}}$ of Gardes and Girard [17], bottom left: estimator $\widehat{\gamma}^{\mathrm{GGS}}$ of Goegebeur *et al.* [20] with $\omega_x = q(u_x|x)$, bottom right: estimator $\widehat{\gamma}^{\mathrm{GGS}}$ of Goegebeur *et al.* [20] with $\omega_x = \widehat{q}(u_x|x, h_x)$.

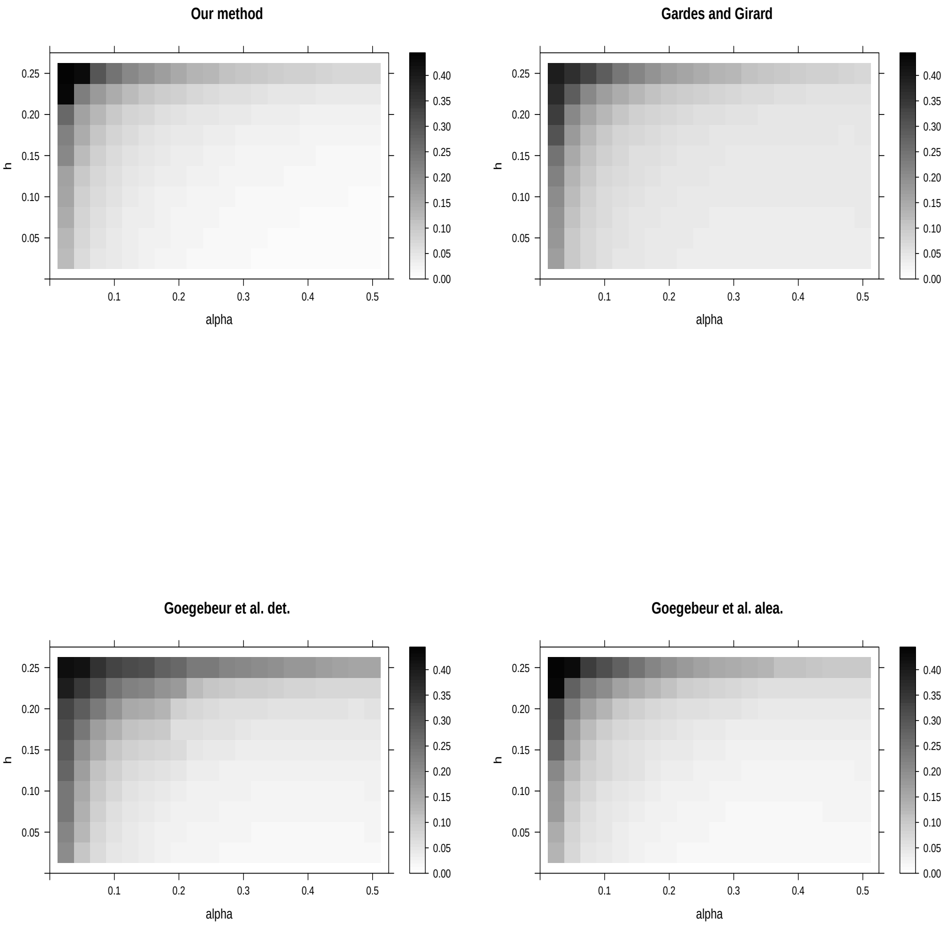

Figure 4: Comparison of the MSE as a function of u_x and h_x for the Fréchet distribution at $x = 0.5$. Top left: our estimator $\widehat{\gamma}$, top right: estimator $\widehat{\gamma}^{\mathrm{GG}}$ of Gardes and Girard [17], bottom left: estimator $\widehat{\gamma}^{\mathrm{GGS}}$ of Goegebeur *et al.* [20] with $\omega_x = q(u_x|x)$, bottom right: estimator $\widehat{\gamma}^{\mathrm{GGS}}$ of Goegebeur *et al.* [20] with $\omega_x = \widehat{q}(u_x|x, h_x)$.

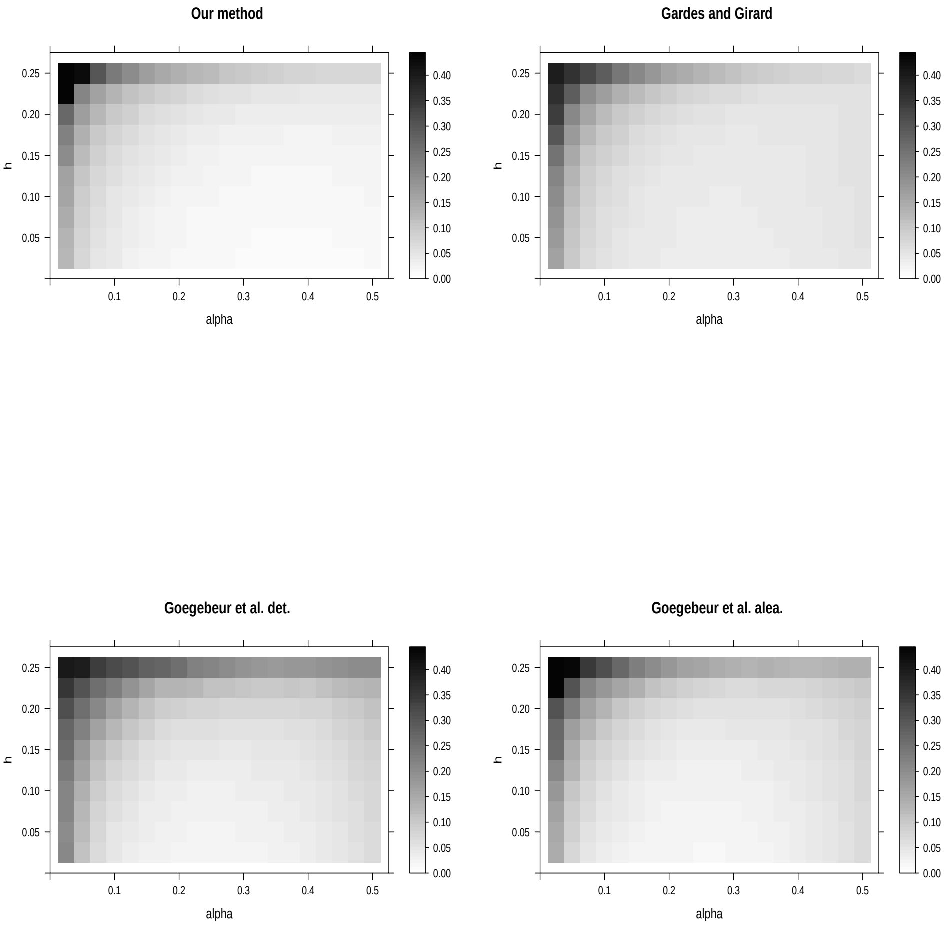

Figure 5: Comparison of the MSE as a function of u_x and h_x for the Student distribution at $x = 0.5$. Top left: our estimator $\widehat{\gamma}$, top right: estimator $\widehat{\gamma}^{\mathrm{GG}}$ of Gardes and Girard [17], bottom left: estimator $\widehat{\gamma}^{\mathrm{GGS}}$ of Goegebeur *et al.* [20] with $\omega_x = q(u_x|x)$, bottom right: estimator $\widehat{\gamma}^{\mathrm{GGS}}$ of Goegebeur *et al.* [20] with $\omega_x = \widehat{q}(u_x|x, h_x)$.

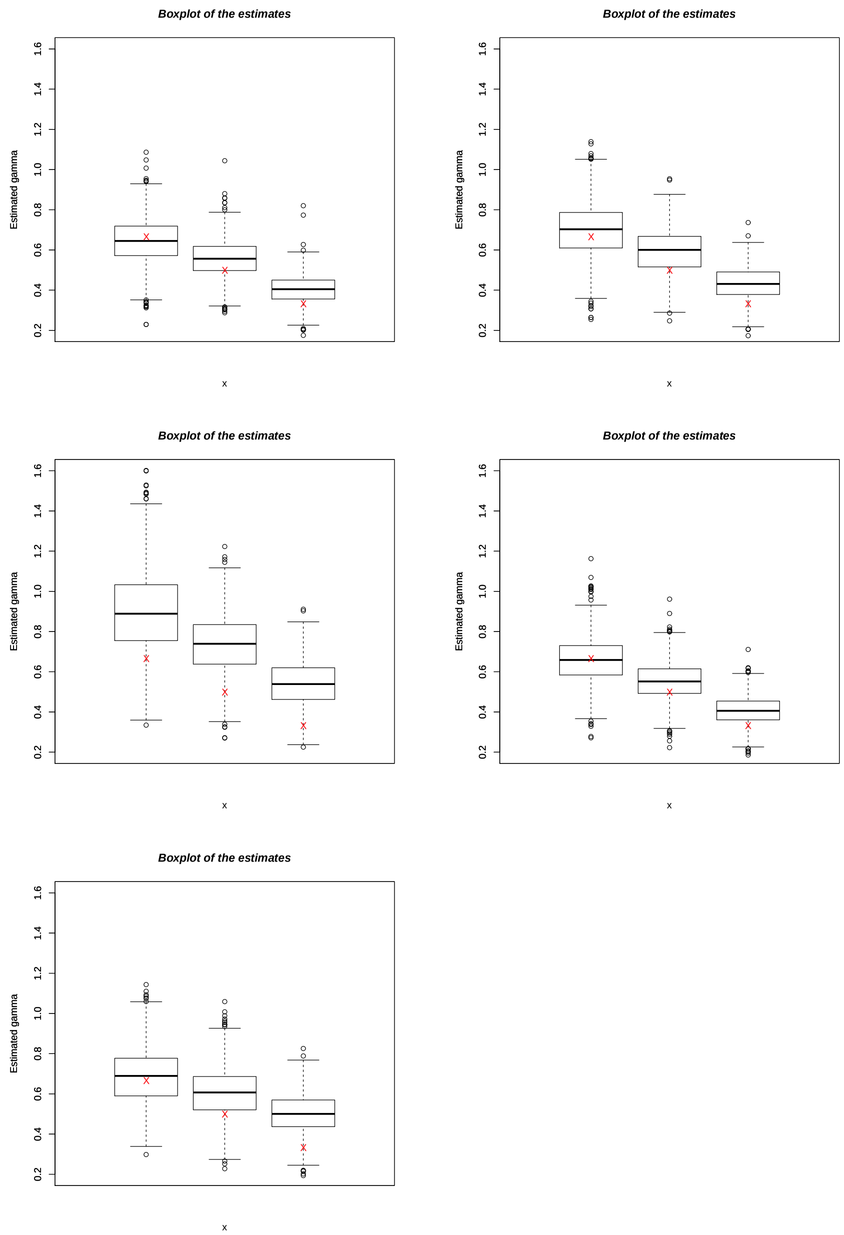

Figure 6: Boxplots of the tail index estimator $\widehat{\gamma}$ for a Burr distribution with $\rho = -3/2$ (top left), $\rho = -1$ (top right), $\rho = -1/2$ (middle left), the Fréchet distribution (middle right) and the Student distribution (bottom left). The red cross is the true value of $\gamma(x)$.

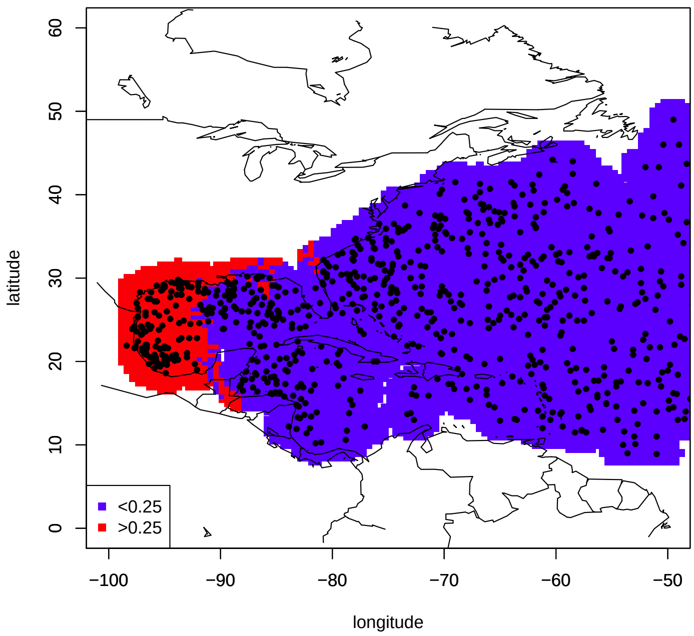

Figure 7: Local estimates of $\gamma(x)$ in the North Atlantic Region. The black dots are the observed locations.